Insuring against Health Shocks: Health Insurance and Household Choices

Kai Liu*

December 2015

Abstract

This paper provides empirical evidence on the role of public health insurance in mitigating adverse outcomes associated with health shocks. Exploiting the rollout of a universal health insurance program in rural China, I find that total household income and consumption are fully insured against health shocks even without access to health insurance. Household labor supply is an important insurance mechanism against health shocks. Access to health insurance helps households to maintain investment in children's human capital during negative health shocks, which suggests that one benefit of health insurance could arise from reducing the use of costly smoothing mechanisms.

*Faculty of Economics, University of Cambridge, Sidgwick Avenue, Cambridge, CB3 9DD, UK, and IZA. Email: kai.liu@econ.cam.ac.uk. I thank the editor and two anonymous referees whose comments improved the paper. I also would like to thank Arne Bigsten, Erzo Luttmer, Magne Mogstad, Marcos Vera and the participants at several seminars and conferences for helpful comments and discussions. All remaining errors are mine.

1 Introduction

Health risk is one of the most important and common risks facing members of a household. The impacts of health shocks may be particularly severe for low-income households, because the cost of treatment can be high relative to income, and the shock may have persistent negative impacts on health given low-quality treatment or a lack of treatment altogether. At the same time, health insurance coverage can be low or even nonexistent for low-income households, particularly in developing countries where formal insurance and credit markets are less developed.

The goal of this paper is to understand the role of public health insurance in mitigating the adverse outcomes associated with health shocks for low-income households. This question is particularly important from a policy viewpoint. For instance, one policy implication from existing research is that the welfare gain from additional social insurance (such as health insurance) may be small, as it would only crowd out other channels for consumption smoothing (Morduch, 1995).[1] However, the welfare gain of health insurance should also depend on its cost relative to other smoothing mechanisms. To infer the welfare value of health insurance, it is then important to examine the efficiency costs of the behaviors used by households to smooth consumption. For example, if a household reduces educational investment in children in order to smooth consumption against health shocks, the value of health insurance could be substantial.[2] To date, there is very little convincing evidence on how access to health insurance interacts with household choices being used as smoothing mechanisms in the absence of health insurance.[3]

In this paper, I analyze the effects of health shocks on income, (food) consumption, and household behavior in terms of human capital and agricultural investments. I estimate the effects of health shocks on these outcome variables when households have access to publicly provided health insurance and compare them to the estimates obtained from households without access to public health insurance. The differences between these two sets of estimates yield the effects of health insurance in mitigating the outcomes associated with health shocks. The main identification problem with a comparison between households covered by health insurance and those without is that

[1] Many studies have emphasized important roles of insurance as provided by household behaviors and institutions, such as savings and asset accumulation (Jalan and Ravallion, 1999), access to microfinance institutions (Gertler, Levine, and Moretti, 2009; Islam and Maitra, 2012), sales of durables and livestock (Rosenzweig and Wolpin, 1993), and family networks (Angelucci, De Giorgi, and Rasul, 2012).

[2] The theoretical frameworks developed in Chetty and Looney (2006) and Chetty (2006) show that when households are very risk averse, they may choose highly costly measures to smooth their consumption paths even though consumption may not fluctuate with the shocks much at first.

[3] In a different but related context, Chetty and Looney (2007) investigate how unemployment insurance provision interacts with household behavior following a job loss by comparing the effects of unemployment shocks on consumption and investments in Indonesia and the US.

insurance coverage is endogenous. For example, individuals with certain unobserved traits (such as risk aversion) may self-select into insurance programs and demand more insurance, and these unobserved characteristics often affect their choices.

To overcome this challenge in identification, I exploit a policy reform associated with the introduction of a large-scale health insurance program in rural China. The New Cooperative Medical Scheme (NCMS) increased the health insurance coverage of rural households from less than 15% before 2000 to over 90% by 2009. The program was implemented over a six-year period from 2003 to 2008 in different counties in rural China at different times. As a result, some areas received coverage earlier than others did, and the households in those areas were then subject to exogenous changes in health insurance status at different points in time. Compared with health insurance reforms in other countries, the Chinese health insurance expansion is attractive for this type of analysis in several ways. First, household mobility in China is restricted because of the household registration system, reducing the extent of any endogenous selection of health insurance arising from mobility. Second, prior to the new program, having health insurance (both private and public health insurance) was quite uncommon among rural households. After the program, the participation rate was very high. Therefore, given the large share of program compliers, the estimated effects lie close to the average treatment effects of health insurance coverage for the general population. Lastly, the level of implementation of the new health insurance program is at the county level, which, combined with limited household mobility, implies that any potential spillover effects will be small.

The estimation is conducted using panel data from the China Health and Nutrition Survey (CHNS). My estimation strategy exploits variations in the timing of the introduction of the program across counties, in a difference-in-difference setting. The identifying variations that I exploit are changes in the effect of a health shock following local introduction of the reform: I compare households' reactions to health shocks before and after the health insurance reform, and between counties that have already implemented the reform and those that have not yet implemented the reform. The identifying assumption is that households living in areas which implemented the reform late and those in areas that implemented the reform early would have had the same time trend in the impacts of the health shock (but are allowed to have different initial levels) in the absence of the program's implementation.

The key findings are as follows. Prior to the reform, a negative health shock is associated with a differential decrease in earnings among household members suffering from the shock. Total

household income (even after excluding transfer income) and household food consumption are fully insured against the health shock. Taken together, this evidence suggests that household labor supply is an important insurance mechanism against health shocks. Given that households were perfectly insured without access to health insurance, one might be tempted to conclude that the welfare value of additional health insurance is small. However, the second set of findings reveals that the benefit of health insurance could come from reducing the use of certain costly smoothing mechanisms. During a negative health shock when there is no access to health insurance, households reduce investment in children's education and agriculture-related activities, and increase the use of child employment. Access to health insurance eliminates the use of some of these costly smoothing mechanisms. Consequently, following a health shock, households with access to health insurance invest more in children's human capital and reduce the use of child employment, relative to the levels that they would have been in the absence of the reform. Given that access to health insurance lessens the financial burden by reducing out-of-pocket medical expenditure and the productivity loss following a health shock, the availability of health insurance mitigates some of the most costly choices households used to make following a health shock.

This paper is connected to a few strands of literature. It is related to a large empirical literature estimating the impacts of health insurance. Recently, much progress has been made using specific policy reforms to evaluate the effects of health insurance on health outcomes and health-related expenditures.[4] Fewer studies have focused on the impact of health insurance on non-health outcomes.[5] Among those that do, few papers have examined the role of health insurance in mitigating adverse outcomes and altering the household choices associated with health shocks.

This paper also contributes to our understanding of how households in developing countries respond to unexpected changes in household resources. Closely related to the current paper is a set of papers analyzing the effects of health shocks on income, consumption and household choices (e.g., Townsend (1994), Kochar (1995), Gertler and Gruber (2002), and Mohanan (2013)). By focusing on a household-level shock and human capital investments, this paper also relates to studies of

[4]See Currie and Madrian (1999) for a review. For instance, a recent paper by Finkelstein, Taubman, Wright, Bernstein, Gruber, Newhouse, Allen, Baicker, et al. (2012) uses the Oregon Medicaid lottery as a source of convincing random variation in health insurance coverage for a group of uninsured low-income adults. A few studies have evaluated the impacts of the NCMS program primarily on health-related outcomes (Lei and Lin (2009), Wagstaff, Lindelow, Jun, Ling, and Juncheng (2009), and Chen and Jin (2012)). Using a different (and cross-sectional) data from 2006 and a matching method, Chen and Jin (2012) also finds some evidence that the NCMS improves school enrolment among six-year-olds.

[5]For example, Gruber and Yelowitz (1999) finds that expansion in Medicaid leads to a reduction in savings and an increase in non-medical expenditures. Similar findings have been found using data from developing countries, such as Chou, Liu, and Hammitt (2003); Wagstaff and Pradhan (2005); Bai and Wu (2014).

the impact of household-level shocks on child labor and schooling (e.g., Beegle, Dehejia, and Gatti (2006) and Yang (2008)). The main distinguishing feature of this study is its use of a novel source of variation in access to health insurance, which makes it possible to examine changes in households' responses to health shocks when access to public health insurance becomes available.

The paper also relates to a small but important literature on whether and how public programs interact with different private channels of insurance. Some studies have found that public health insurance crowds out the purchase of private insurance (Cutler and Gruber, 1996) and increases the receipt of certain public transfer programs (Baicker, Finkelstein, Song, and Taubman, 2013). Cox, Eser, and Jimenez (1998), Attanasio and Rıos-Rull (2000) and Jensen (2004) conclude that an increase in the benefits from public transfer programs crowds out private transfers which were used to support extended family members. The results of the current paper add to these findings by showing that health insurance may crowd out other private arrangements to cope with health shocks, some of which may be quite costly and could have very different welfare implications of a public health insurance program.

The paper proceeds as follows. Section 2 discusses the institutional background of the health insurance reform. Section 3 presents the empirical model, followed by a discussion of the data and the main variables used in the estimation in Section 4. The estimation results on income and expenditure are presented in Section 5. The estimation results on investment are presented in Section 6. Section 7 conducts robustness and specification checks. Section 8 concludes.

2 Institutional Background: The Chinese Health Insurance Reform

Prior to the transition from a planned economy to a market economy, households in rural China had access to universal health insurance through the Cooperative Medical Scheme (CMS). When China reformed its rural economy in 1979 and introduced the Household Responsibility System, the CMS collapsed (because of a lack of public funding) and left around 90% of all farmers uninsured (Yip and Hsiao, 2008). In the next two decades, households in rural China had little formal insurance against health shocks. During the 1990s, despite several attempts to rebuild the CMS, the coverage rate remained below 15%.

Aiming to provide a basic social safety net against the cost of illness for *all* rural households, in 2003, the central government launched the NCMS with the goal of offering health insurance to

all rural areas by 2010. The new program eventually replaced the old CMS program. The introduction of the new program was rolled out gradually at the county level. Provincial governments, following guidelines from the central government, determined the timing of the establishment. In 2003, each provincial government was required to select at least 2–3 counties within the province as pilot counties for the NCMS, with the selection of the pilot counties based on three broad criteria: a willingness to participate among the rural residents, fiscal soundness, and a solid foundation for management (State Council, 2003). In 2006, the central government required provincial governments to expand the program to include at least 40% of all counties by the end of 2006 and 60% by 2007 (Department of Health, 2006). By 2008, the program had been implemented in all rural areas covering the entire rural Chinese population. Therefore, for nearly half a decade, rural households in China experienced different access to public health insurance, and whether one was covered depended on the county in which the household was registered. Because of the strict household registration system (*hukou*), mobility of households between counties is restricted.

Participation in the program is voluntary. As a general rule, if the household decides to join the program, all members of the household must enroll in the program. The annual premium is kept low thanks to heavy subsidies from the government. For example, in 2008, the typical annual premium was 20 RMB (3 USD) per person, supplemented by a subsidy of 80 RMB (12 USD) from the central and local governments (Department of Health, 2008).[6] The amount of subsidies has increased over time, and the central government provides larger subsidies for underdeveloped regions in the central and western parts of China. In the first few years of the reform, the program emphasized the coverage of inpatient expenses and outpatient expenses related to severe illness.[7] In recent years, there have been efforts to increase the coverage of outpatient expenses and to decrease the deductables for inpatient services (Department of Health, 2008). The benefits of the program vary by county. Two-thirds of the counties cover both inpatient service and outpatient service, with the cost of inpatient service reimbursed through a formula and the cost of outpatient service reimbursed through a household medical savings account.

Table 1 summarizes the expansion of the health insurance program in the sample used in this analysis.[8] Among the 36 counties observed in the data, three had commenced the program by 2004. The program expanded quickly, with over 50% of counties having implemented the program

[6] In 2006, about 20% of the costs of the program were covered by the central government, 50% was financed by the local government, and the remaining 30% was from the households' premium payments.

[7] See Lei and Lin (2009) for more on the insurance coverage. Wagstaff, Lindelow, Jun, Ling, and Juncheng (2009) provides a detailed description of the insurance coverage for a sample of pilot counties in 2005.

[8] Details of the data and sample construction are discussed in Section 4.

by 2006. By 2009, all the counties in the data were covered by the NCMS. Table 2 reports the same pattern of increase among households with access to the NCMS. The proportion of households that are insured by either the old CMS or the NCMS grew quickly after the implementation of the reform, rising from 7.0% in 2000 to over 90% in 2009. As the NCMS had replaced the old CMS in all counties in 2009, we could infer that the take-up rate was 93.5% in 2009. The last two columns in Table 2 show the percentage of households covered by any kind of health insurance, private and public combined. Health insurance coverage never exceeded 15% before the first year of the reform (2004), but began to increase steadily thereafter with the introduction of the NCMS. This pattern is confirmed graphically in Figure 1, which plots the average insurance coverage over time (in terms of survey waves) relative to the health insurance reform. For each county, I recenter the data such that time zero is equal to the first wave observed since the reform. The picture shows a dramatic and persistent increase in health insurance coverage in rural China associated with the reform.

It is useful to understand the determinants of the timing of the reform across the sample counties. For this purpose, I regress the year of implementation on a range of observed county characteristics measured before the reform, including average number of days of severe sickness, average log household income, average proportion of households with any health insurance, average level of education and age of the household head. Table 3 reports that none of the estimated coefficients is significantly different from zero at the 10% level. Overall, in the sample, there is little evidence of a relationship between most observable factors and the timing of the implementation of the reform. Note that, as discussed in the estimation strategy in the following section, any time-invariant characteristics at the county level are controlled for using county fixed effects. Therefore, it is not necessary that the timing of the reform be unrelated to county-specific characteristics.

3 Empirical Strategy

In estimating the impact of health shocks on changes in the outcome variables between survey waves, I use the following first-differenced regression specification (e.g., Gertler and Gruber (2002)):

$$\triangle y_{ijt} = \lambda \triangle h_{ijt} + \beta X_{ijt} + \gamma_j \times \gamma_t + \varepsilon_{ijt} \tag{1}$$

where $\triangle y_{ijt}$ is the change in an outcome of interest for household i residing in county j at year t, and $\triangle h_{ijt}$ is the health shock for household i (to be defined in Section 4). The first-difference setup is equivalent to the inclusion of household fixed effects in a levels regression; it already

eliminates unobserved household heterogeneity (such as preferences and health endowment) which may determine the outcome variable and illness at the same time. X_{ijt} includes a set of demographic variables, including change in log household size, change in the share of children in the household, change in the share of working-age adults in the household, age and age-squared of the household head, indicators for the education level of the household head, and whether the household head is single. The inclusion of X_{ijt} controls for changes in the outcome variables related to household demographics. γ_j and γ_t are county and year fixed effects, respectively. $\gamma_j \times \gamma_t$ are county-by-year fixed effects, which control for the average change in outcomes across all households or the change in aggregate resources within a county in a given year.[9]

The coefficient of interest is λ, the impact of a unit change in the health shock on the outcome variable. If the outcome variable is the change in household consumption, equation (1) constitutes the standard test for consumption smoothing used in the literature (e.g., Cochrane (1991), Mace (1991), Townsend (1994)). Under the null hypothesis of complete insurance, the coefficient on health shock, λ, should be zero. If households are fully insured against idiosyncratic shocks (through either mutual insurance or self-insurance), consumption growth should not be correlated with changes in health once growth in county-level resources is controlled for.

To estimate household behavior following a health shock with and without access to health insurance, I extend the model in equation (1) by allowing the coefficient λ to vary with the reform:

$$\lambda_{ijt} = \alpha_0 + \alpha_1 R_{jt} + \gamma_j + \gamma_t + \nu_{ijt} \tag{2}$$

where R_{jt} is a dummy variable which equals one if the health insurance reform has been implemented in county j by year t. The identifying variations that I exploit in this specification are county-specific changes in the coefficient on the health shock after local introduction of the reform. The year fixed effects (γ_t) allow for secular changes in λ over time that may be completely unrelated to the reform. The county fixed effects (γ_j) allow for the fact that variation in the timing of the reform across counties may not have been exogenous. Consistent estimation of α_1 is obtained so long as (a) these characteristics are fixed over time during the sampling periods or (b) implementation of the reform is not correlated with changes in these characteristics.[10]

[9] For example, this may capture local economic growth and potential changes in the local strictness of mandatory schooling.

[10] Note that the key parameters of interest is how λ varies before and after the reform. Therefore, these identifying assumptions are made with respect to a marginal change in $\triangle h_{ijt}$. For instance, suppose that the implementation of the NCMS is perfectly correlated with local changes in the strictness of the mandatory schooling law (which is allowed for under the model–see more on this below). The second assumption says that any changes in the local

Substituting equation (2) into equation (1), I obtain the following main empirical model:

$$\triangle y_{ijt} = \alpha_0 \triangle h_{ijt} + \alpha_1 (\triangle h_{ijt} \times R_{jt}) + \triangle h_{ijt} \gamma_t + \triangle h_{ijt} \gamma_j + \gamma_j \times \gamma_t + \beta X_{ijt} + \varepsilon_{ijt} \tag{3}$$

$\triangle h_{ijt} \gamma_t$ are wave fixed effects interacted with health shocks, which allow for a flexible time trend in the effects of health changes (e.g., from country-wide changes in health care). $\triangle h_{ijt} \gamma_j$ are county dummies interacted with health shocks, which control for persistent differences among counties in the severity of health shocks or the quality of health care. To capture any common shock to the outcome variable across households within the county and over time, standard errors are clustered at the county level.[11]

Coefficient α_0 measures the direct effect of a health shock in the absence of health insurance.[12] α_1 identifies the effect of health insurance coverage in mitigating the adverse outcomes associated with health shocks. In the context of consumption insurance, it is the effect of health insurance on consumption smoothing against health shocks. The predicted effect of a health shock after the health insurance reform is given by the sum of α_0 and α_1.

While the first-difference setup in equation (1) eliminates permanent unobserved household heterogeneity, the estimated parameter α_0 may still be biased because of unobserved differences in the changes in outcomes across different types of households. What are examples of situations that would violate my identifying assumption? Changes in aggregate resources, such as the supply of health care and local economic conditions, may affect the health shock and changes in outcomes. However, the inclusion of county-year fixed effects already controls for unobserved differences across counties *and* over time. Household mobility is restricted, which limits the extent of self-selection into areas with better health services through regional mobility. Perhaps a more plausible source of bias is unobserved shocks to household resources which induce the health shock. For instance, previous job displacement may simultaneously affect current health (and hence a change in health since displacement) and changes in outcomes (Black, Devereux, and Salvanes, 2012). In Section 7, I show that there is little correlation between past income shocks and the current health shock, indicating that this potential source of bias may be less of a concern in my sample. Note that even

implementation of mandatory schooling law does not affect children from households suffering from a health shock and children from households without a health shock differently.

[11] Note that, because of the inclusion of county-by-year fixed effects, the reform indicator R_{jt} is excluded from the regression. Therefore, the direct effect of being covered by health insurance (on households that did not experience health shocks) is not identified using this framework.

[12] In all of the tables, this coefficient is the overall mean effect of a health shock in the absence of health insurance, as opposed to the effect in the base county and base year.

if the effect of the health shock is contaminated by the effect of other types of shocks to household resources, the estimated coefficient α_1 will be unbiased as long as the omitted variable bias on the coefficient of $\triangle h_{ijt}$ does not change with the reform.[13] This is more of a concern if the reform affects the unobserved composition of households reporting any illness. In Section 7, I provide some evidence that the reform is unlikely to affect the distribution of illness across households.

When the outcome variable is the change in household consumption, the marginal utility of consumption may depend on health status directly, or indirectly through induced changes in leisure when consumption and leisure are not additively separable. In this context, the health shock will be endogenous, as it is correlated with omitted preferences in the error term. As a result, the growth in consumption will vary with the state of health, even with full insurance. I investigate the importance of this bias in Section 5.2, by drawing upon individual consumption data from household members not experiencing a health shock.

4 Data, Sample Construction and Variable Definitions

I use data from the CHNS. The survey is based on a multistage, random cluster process that yields a sample of about 4,400 households with 19,000 individuals tracked over time. The survey contains individual-level information on health, income, (food) consumption, education and other demographics and household information. The sample covers nine provinces which vary substantially in terms of geography, economic development, and other socioeconomic indicators. The survey was conducted in each of the years 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009, and 2011.

I use data from 1993–2011, comprising of seven waves covering the entire duration of the health insurance reform. Because the outcome variables and the health shock are in first differences, the data from 1993 are used as the initial condition.[14] Given that enrollment in the health insurance program is at the household level, the unit of observation under analysis is the household. The sample consists of rural households, defined as those in which the household head lives in a rural county and has a rural *hukou*.[15] Given that a health shock is defined in terms of the changes in health between waves among key members of the household (as explained below), I focus on stable households that have not experienced any change in household head.[16] In order to capture health

[13] Formally, let Ω_{it} be all covariates in the regression (3), excluding R_{jt}. The estimated coefficient α_1 is unbiased if $E(\triangle h_{ijt}\varepsilon_{ijt}|R_{jt}, \Omega_{it}) = E(\triangle h_{ijt}\varepsilon_{ijt}|\Omega_{it})$.

[14] After taking first differences, the sample consists of the six waves collected in 1997, 2000, 2004, 2006, 2009, and 2011.

[15] A rural *hukou* is required in order to enroll in the NCMS.

[16] In the sample, the head of household is typically a male, the primary earner in the household, and the oldest

shocks with a significant economic impact, I exclude households where the head of the household is more than 65 years of age.

Following Lei and Lin (2009), I draw from the community survey data from the CHNS to determine the year of implementation of the NCMS at the county level.[17] In this survey, government officials from each community were required to respond whether the CMS had been implemented in their community, and if so, the starting date. As the pilot implementation of the NCMS started in 2003, it is clear that those communities that began the CMS before 2003 were implementing the old CMS. Therefore, CMS plans that started to operate in or after 2003 are defined as the NCMS. Given that the NCMS implementation takes place at the county level, if any community within a county was known to have adopted the NCMS, then the county as a whole was defined as having implemented the NCMS.[18]

The measure of consumption available from the CHNS is food consumption. Food consumption accounts for roughly half of total household expenditure in rural China during the sample period (Yu and Abler, 2009). These data are collected by a dietary questionnaire recording the items and amounts of food consumption for each member of the household over the past three days (including food consumed both at and away from home). The CHNS calculates macronutrients based on the individual food consumption data. The macronutrients include the average daily intake of calories (in kilocalorie), protein (in grams), fat (in grams) and carbohydrate (in grams). Monetary value of the food consumption is, however, not directly available from the data. I trim the top and bottom 0.25% of the daily caloric intake (and the other nutrients associated with the outliers in calories), corresponding to daily caloric intakes that are above 4,859 kilocalories and below 433 kilocalories. Household consumption is defined as the average macronutrient intake among adult individuals (18 years or older).[19] One advantage of working with nutrition intake (as opposed to the monetary value) is that one does not need to compute the value of home-produced food, as the prices of home-produced food are usually difficult to predict. In addition, while the budget share of food consumption decreases with income, nutrition intakes such as protein are luxury goods which could

person in the household. About 15% of households experienced a change in household head and so were excluded from the sample.

[17] While the household survey has a question relating to whether an individual had enrolled in any Cooperative Medical Scheme (i.e., the CMS), prior to the 2009 survey, it does not distinguish between the old CMS and the NCMS.

[18] Note that the type of benefits package in the NCMS is not available in the data for any given county.

[19] I also construct household consumption by dividing total food consumption (among all members) by an adult equivalence scale. The adult equivalence scale used is one for household members aged 18 years or older, and 0.73 otherwise (Angelucci, De Giorgi, and Rasul, 2012). The results are robust to this alternative measure of household consumption.

be a better proxy of nondurable consumption over time (Angelucci and Attanasio, 2013).

The health shock measure. The derivation of the variables related to health status is from two questions in the survey. The first asks whether the individual, during the past four weeks, had been sick, injured, or suffering from a chronic or acute disease (*Any Illness*); if the response is yes, the second question records the number of days that the person had been unable to carry out daily activities due to the illness in the same period (*Severe Illness*).[20] The second question is intended to capture only severe illness, both chronic and acute, that damages in the ability of to perform daily activities. Both questions are available in all waves of the survey for every respondent aged 18 years or older.[21]

I construct a household-level health shock variable as follows. Let d^h_{ijt} and d^s_{ijt} indicate the number of days of sickness in the past four weeks for the household head and the spouse of the household head, respectively. The health status for household i living in county j in year t is defined as the sum of d^h_{ijt} and d^s_{ijt}, normalized by the length of the recall period (28 days):

$$h_{ijt} = \frac{d^h_{ijt} + d^s_{ijt}}{28} \times 100 \tag{4}$$

Consequently, the health shock is defined as the change in h_{ijt} between waves.[22] h_{ijt} can be interpreted as the proportion of time in the past four weeks (in percentage points) that the household suffers from a severe illness.

The above measure of health shock focuses on changes in the health status of the head and the spouse of the head. I also consider an alternative measure of health status by taking into account of the number of days during which other household members reported severe illness:

$$h_{ijt} = \frac{d^h_{ijt} + d^s_{ijt} + \sum_{m=1}^{M} d^m_{ijt}\mathbf{1}(18 \leq age_m \leq 55)}{28} \times 100 \tag{5}$$

[20] The survey questions are: "During the past four weeks, have you been sick or injured? Have you suffered from a chronic or acute disease?" and "For how many days during the past four weeks were you unable to carry out normal activities due to this illness?".

[21] Another potential measure of health shocks often used in the literature is based on the self-reported ability to perform daily activities (Strauss, Gertler, Rahman, and Fox, 1993; Gertler and Gruber, 2002). This is not pursued here because the CHNS only collects data on physical limitations for persons aged over 55 years, and the collection of these data ends with the 2006 wave. This means that the health shocks defined by physical limitations are only available for the elderly population, and health shocks are not defined for the half of the counties that implemented the reform after 2006.

[22] For an individual without any sickness in the past four weeks (i.e. responds "no" to the first question), the number of days of sickness is coded to zero. Note that it is possible to have the number of days limiting daily activities equal to zero when the individual answers "yes" to the first question (e.g., minor illness not affecting daily activities).

where the nominator sums over the number of days of severe illness among the remaining household members aged between 18 and 55 years (in addition to the head and the spouse of the head).

The constructed health shock is a mixture (or weighted average) of permanent and transitory health shocks. To investigate the nature of the health shock, in Appendix Table A1, I report results from a regression of the current health shock (defined for the head and the spouse of the head) on the first- and second-lagged health shock. After controlling for county-by-year fixed effects, the autocovariance structure between the current health shock and the first- and second-lagged health shocks are -0.66 and -0.31, respectively.[23] This indicates a strong mean-reverting pattern for health shocks, indicating that the health shocks in the paper are far less permanent in nature. For health shocks that are more persistent, the estimated effects of the health shock and the associated role of health insurance in mitigating these negative effects are likely to be a lower bound (in absolute values).[24]

One limitation regarding the measurement of health shock in the data is that the recall period of the respondent's health condition is four weeks. Certain outcome variables (such as income and investments in agriculture) are based on values over the past 12 months. The difference in recall period may imply that the outcome variable may appear "excessively smooth" to the health shock (coefficient α_0 biased toward zero), especially if the impact of the health shock in the past four weeks is small relative to the level of outcome prior to the health shock. The assumption underlying the regressions when outcome variables are measured annually is that the percentage of days with severe illness in the past four weeks is a proxy for the percentage of days with severe illness over the past year. Note that, even in cases where the estimated coefficient α_0 is biased toward zero, in a difference-in-difference setting, the estimated effect of the reform in mitigating health shocks (coefficient α_1) remains unbiased. As a robustness check, in Section 7, I conduct the analysis using changes in the number of weeks of sickness affecting work and daily activities over the past three months as a measure of health shock. The estimates are qualitatively similar to the estimates from the baseline model.

[23]The county-by-year fixed effects are included such that the health shock is defined in a similar fashion as the main regression. The coefficients on the e lagged health shock terms are the estimated autocovariances of health shock at e lags.

[24]Identifying the impact of health insurance in mitigating negative impacts from permanent health shocks and from transitory health shocks is, however, difficult given the nature of the data. For instance, one category of a permanent health shock is disability. The CHNS data contains several outcomes from physical examinations from which I can construct a measure of disability for each individual (an individual is disabled if he is blind, or suffering loss of at least one arm and/or at least one leg). Using this definition, on average only 0.62% of the households in the sample are disabled in a given year (i.e. either the head or the spouse of the head is disabled). The sample size of the disabled households is not sufficient to undertake the analysis.

Attrition in the sample. Because the health shock and other outcome variables in analysis are defined using changes in variables between adjacent waves of the survey, it is important to consider attrition within the selected sample. At the household level, about 8.3% on average cannot be followed between the survey waves. This is a low attrition rate for a panel survey, particularly in a developing country. Attrition is potentially worrisome if it is correlated with the independent variable of interest, such that sample selection could then lead to biased estimates. I estimate regressions of equation (3) on the overall sample used in the main analysis. The dependent variable is an indicator variable equal to 1 if the household cannot be tracked in the following survey round (and 0 otherwise). There is no evidence that attrition is correlated with the health shock or the interaction term between the health shock and the reform: the coefficient estimates α_0 and α_1 are small in magnitude and statistically insignificantly different from zero.[25]

At the individual level, on the other hand, sample attrition is higher especially among children. The main outcome variables on children used in the analysis are school attendance and work status. For these variables, I explain the tracking of household children across survey rounds in Appendix B.

Summary statistics. Table 4 reports summary statistics for the households used in the empirical analysis. The average earnings of the head and spouse of the household head represent 67% of household total income, with another 10% of average household income being from transfers and subsidies. Decomposing total household income by the type of activities, I find that about 40% of average household income is from agricultural-related activities, with around 80% of the sample households engaging in farming and about half in raising livestock. For either health shock measure, the average share of days with a severe illness within a household is a little over 2 percentage points, with one standard deviation of about 12 percentage points.

5 Estimation Results: Income and Expenditure

This section examines the impact of household health shocks on the following outcomes in sequence: household income and its components, food consumption and medical expenditure.[26] In Section 5.1, I begin by estimating the direct effect of a health shock on productivity, using earnings among household members reporting severe illness as the outcome variable. I then examine the effects of

[25]The estimated coefficient α_0 is 0.0003 with a standard error 0.0008; the estimated coefficient α_1 is -0.0002 with a standard error of 0.0008.

[26]Definitions of these outcome variables are described in Appendix B.

a health shock on household income and its components. This is important because a health shock affects specific members of the household, and for rural households, there is large potential for income smoothing by family labor supply and the household's choosing certain economic activities (Morduch, 1995). The coefficient estimate on the interaction of health shocks and the reform shows the role of health insurance in mitigating the effects of health shocks. Having obtained estimates on household resources, in Section 5.2, I estimate the effects of health insurance on consumption smoothing and medical expenditure related to treatment of the illness.

5.1 Household income and its components

Table 5 presents the coefficient estimates on the health shock and the interaction of the health shock and the reform. For each definition of health shock, the outcome variables are the changes in the log earnings (among a subset of household members for which the health shock is defined), log household income, and log household income excluding subsidies and transfer income. When the health shock is defined on the household head and the spouse of the head, I also investigate the impact on log income from other household members. Panel (a) reports the estimates when the health shock is defined as the percentage of the past four weeks during which the head and spouse reported severe illness. The coefficient on the health shock (α_0) indicates that, prior to the reform, a 10-percentage-point increase in the size of the health shock is associated with a differential decrease in the total income of the head and the spouse of the head of 17%. After the reform, the negative impact on income reduces by one-half. The difference (α_1) is only borderline significant. Therefore, without health insurance, illness facing the head and the spouse of the head had a significant negative impact on their productivity. There is some evidence that health insurance dampens the productivity losses arising from a health shock experienced by the head and the spouse of the head, supporting the hypothesis that health insurance potentially improves the quality of treatment for illness.

The coefficients on the health shock in the regressions for total household income are close to zero and not statistically significantly different from zero (second row of Panel (a)). The coefficient on the interaction of the health shock and the reform is also insignificantly different from zero. Contrary to the direct productivity loss arising from the health shock, total household income appears to be well insured against health shocks, even without access to health insurance. One might think that transfer income and subsidies play an important role. However, the coefficient on the health shock in the regressions for total household income excluding subsidies and transfers

are also close to zero and not statistically significantly different from zero (third row of Panel (a)). This suggests that transfers and subsidies play a minimal role in mitigating the negative impact of a severe health shock. Taken together, the impact of the shock on total income appears to work entirely through the change in the income from other household members: the last row in Panel (a) presents evidence that income from other household members compensates for the productivity losses incurred by the health shock on the head and the spouse of the head. The reform appears to reduce the extent of such compensatory behavior, although the coefficient on the health shock interacted with the reform is not very precise.

Panel (b) presents the coefficient estimates for when the health shock is based on the share of the past four weeks with severe sickness among the head, the spouse of the head, and all other working-aged household members. Prior to the reform, a 10-percentage-point increase in the size of the health shock was associated with a differential decrease in the total income of the head, spouse of head, and all other working-aged members of 4%. The effect is much smaller than the effect of health shock on total earnings from the head and the spouse of the head reported in the first row of Panel (a), which provides further evidence of compensatory labor supply behavior among the primary household members. Similar to the evidence presented in Panel (a), the health shock has a negligible impact on total household income, both with and without transfers and subsidies. This suggests that the labor supply from remaining household members helps to smooth household income following a health shock to the primary household members.

5.2 Household food consumption and medical expenditure

Table 6 reports the coefficient estimates when the outcome variables are the changes in household food consumption across four nutritional intakes: calories, proteins, fats and carbohydrates. If households are fully insured against idiosyncratic shocks, consumption growth should not be correlated with changes in health once growth in county-level resources is controlled for, and consequently, the coefficient on health shock should be zero.

I find that prior to the reform, negative increments to health are not associated with reductions in the intake of macronutrients. The point estimates are small and not statistically different from zero. Therefore, I cannot reject the hypothesis of full insurance against severe health shocks before the reform. This is perhaps not so surprising, given that households were already able to smooth household income against health shocks to specific members.

In Table 6, the coefficient estimates on the interaction of the health shock and reform is the

estimated consumption smoothing effects of the reform (parameter α_1). None of the point estimates are significant. Adding the coefficients in the first and second rows gives the predicted net effects of the health shock after the reform. The t-test reported in the third column implies that after the reform, I still cannot reject the hypothesis that households are fully insured against health shocks.[27]

One potential concern is that consumption growth may be shifted by changes in the state of health (even with full insurance) when the marginal utility of consumption depends on health status, either directly or indirectly through induced changes in leisure if consumption and leisure are not additively separable. One advantage of the CHNS data over other household panel data of consumption is that they collect food consumption data at the individual level. I therefore construct an alternative measure of adult-equivalent consumption by excluding the consumption of the household head and/or the spouse of the head whoever experienced a health shock in the panel. This measure circumvents the problem that changes in health may directly impact the marginal utility of consumption for the sick member. Table 7 presents the estimated coefficients when the outcome variables are the changes in log of consumption, excluding the consumption of any members reporting any illness. With the exception of fat intake, the remaining macronutrient intakes are completely insured against health shocks, both before and after the reform. This indicates that the potential bias arising from state-dependent preferences in consumption is small.

The direct effect of health insurance is to reduce out-of-pocket medical costs to treat severe sickness, which can be substantial relative to the income level. To evaluate the effect of the reform on health expenditures related to treating the reported severe illness over the past four weeks, I use the following regression framework:

$$y_{ijt} = b_0 + b_1 R_{jt} + b_2 \times Sick_{ijt} + \gamma_j + \gamma_t + \varepsilon_{ijt} \tag{6}$$

where, as before, R_{jt} is the reform indicator, and γ_j and γ_t are the county and calendar year fixed effects, respectively. y_{ijt} measures total out-of-pocket medical expenditures among members of household i for which the health shock is defined.[28] $Sick_{ijt}$ is an indicator of whether any member of the household (for which the health shock is defined) suffered from any illness in the past four

[27]The estimated coefficients on the main control variables are not reported but show the expected sign as in the literature. For instance, the growth in per capita log consumption falls with the increase in log household size, indicating some economies of scale in consumption. Consumption grows at a faster rate for household heads who have at least graduated from primary school.

[28]The out-of-pocket cost is the proportion of the cost paid by the individual being treated for the same illness that was used to define the health shock. See Appendix B for details.

weeks, which controls for the mean difference in y_{ijt} between households reporting any sickness and those reporting none.[29]

Table 8 reports the coefficient estimate on the reform indicator. In column (1) of Table 8, I report the estimated coefficient b_1 where the dependent variable is an indicator for having paid positive out-of-pocket medical expenditure. Coefficient b_1 measures whether the likelihood of paying for medical expenditure out-of-pocket for treating the illness changes after the reform. As shown, the reform reduces the incidence of positive health expenditure by about 5 percentage points. The reform does not appear to have any impact on total health expenditure at the intensive margin (i.e., log out-of-pocket expenditure). The estimated coefficient in column (2) is insignificantly different from zero. One potential explanation is that, after the reform, households may switch to better-quality health care, leaving the total out-of-pocket expenditure unchanged.[30] Taking the effects at both the extensive and intensive margin, the Tobit regression estimates in column (3) imply a significant reduction in health expenditure after the reform for treating the illness.

Three conclusions emerge from the analysis thus far. First, prior to the reform, the estimated impact of health shocks on total household income and food consumption is small and insignificantly different from zero. Second, prior to the reform, the effects of a negative health shock were concentrated on the earnings of household members directly affected by the shock and on out-of-pocket medical expenditures made to treat the illness. Taken together with the implications on total household income, this suggests compensatory labor supply behavior against the health shock among other members of the household not experiencing the health shock. Third, after the reform, there is some (weak) evidence that health insurance reduces the productivity loss among members suffered from the health shock. The reform also lessens the financial burden on the household by reducing out-of-pocket medical expenditure associated with treating the illness.

6 Estimation Results: Investment

This section focus on investment choices in response to health shocks and the role of health insurance in changing these choices. As outcomes related to investment choices, I consider human capital investment in children (Section 6.1) and investments related to agricultural activities (Section 6.2).

[29] Health expenditures are only recorded where they are associated with treating the illness that the individual had (both severe and non-severe illness) within the past four weeks. Therefore, by definition, health expenditures should be zero for those not reporting any illness.

[30] Lei and Lin (2009) and Wagstaff, Lindelow, Jun, Ling, and Juncheng (2009) provide some evidence that the NCMS increases the utilization of formal outpatient and inpatient care and reduces the use of traditional folk doctors.

Determining why and how households smooth consumption is important for evaluating the welfare consequences of insurance policies. The availability of public health insurance means that households may switch from using costly smoothing mechanisms to public health insurance in response to health shocks. In the context of Chetty and Looney (2006), the availability of health insurance to cover the lumpy expenditures associated with health care can be thought of as an additional channel to insure against health shocks. If public health insurance is a cheaper means to insure against health shocks, and if households are very risk averse, the household will substitute public health insurance for more costly mechanisms (even when there is perfect consumption smoothing against health shocks).[31]

6.1 Human capital investment in children

Table 9 presents coefficient estimates on the health shock and the interaction between the health shock and the reform when the outcome variables are individual-level changes in school attendance and employment. Individuals were included in the analysis if they were aged 7–18 years.[32] Because childrens outcomes are defined as changes in schooling or changes in employment at the individual child-level, the first-differenced regression model (equation (3)) absorbs child-specific unobserved characteristics. To the extent that migration out of the household (e.g. female daughter leaving home, son getting married) is predicted by the child-specific fixed effect, this specification is robust to potential selection bias caused by changes in the composition of children within a household.[33]

I focus on discussing the estimates from the health shock defined on the head and the spouse of the head, given that the estimates using the alternative definition of health shock are quite similar. The first column of Panel (a) in Table 9 shows that, in the absence of health insurance, a 10-

[31]The model based on Chetty and Looney (2006) is outlined in Appendix A.

[32]In China, compulsory schooling is 9 years. The usual age finishing 9 years of schooling is 16 years. After then, students may choose to attend high school (3 years), to receive vocational training (2-3 years), or to seek employment. Tuition is free for the 9 years of schooling (primary and lower-middle school education), although students normally have to pay costs associated with purchasing books and tools, transportation, fees related to extracurricular activities, and fees for boarding school (which is quite common in rural areas). For instance, Brown and Park (2002) document that in poor counties in rural China, a family with one child in primary school and another in lower middle school spends as much as half of their per capita expenditures on fees related to education. In addition, there are opportunity costs in terms of forgone earnings (especially from employment opportunities in cities), which may be increasingly high as children enter their late teens (Meng, 2012).

[33]I also do not find any significant effect of health insurance access on household structure (e.g. number of school-aged children and share of children within a household). Another potential concern is that the sample of households with school-aged children in two consecutive survey waves (when changes in outcomes are defined) is different from that for previous outcomes on income and consumption. The estimated effects of a health shock on household income and consumption are qualitatively similar on a smaller sample of households with at least one school-aged child in two consecutive surveys (results not reported but available upon request). Note that, in all regressions, I control for household demographic characteristics and changes in household compositions (see Section 3 for details).

percentage-point increase in the size of the health shock (on the head and the spouse of the head) is associated with a differential decrease in the rate of school enrollment of 6 percentage points. Column (2) shows that a 10-percentage-point increase in the size of the health shock is associated with a differential increase in the rate of children at work of 8 percentage points. Therefore, without access to public health insurance, it appears that one way to maintain household income and consumption is to reduce educational expenditures by taking children out of school and increasing the household labor supply through the use of child employment.

The coefficient on the interaction term between the health shock and the reform (α_1) identifies the differential effects of the health shock after the reform. Access to health insurance completely mutes the effects of the health shock on the investment in children's education: a 10-percentage-point increase in the size of the health shock is associated with a differential increase in the rate of school enrollment within a household of 5 percentage points, relative to households experiencing the same health shock without health insurance (second row of coefficients). The net effect of a health shock with health insurance on school enrollment is small and statistically insignificantly different from zero ($\alpha_0 + \alpha_1$). Access to health insurance also reduces the use of child employment as a way of coping with health shock: the coefficient estimate of α_1 is -0.006, which implies that, after the reform, the net effect of a health shock on the use of child employment is insignificantly different from zero (third row in the second column).

Tables 10-11 report heterogeneity impacts of health shocks and health insurance on human capital investments. Panels A of Tables 10-11 exploit the heterogeneity of the health shock and their differential impact on childs schooling and employment. I expand the main regression model (equation (3)) by including two health shocks: one is the health shock to the head (typically male) and the other is the health shock to the spouse of the head.[34] The results show that, holding the health shock to the spouse of the head constant, health shock to the head has a significant negative impact on schooling and positive impact on child employment. The magnitude of the effects is much larger than the estimated effects from the health shock on the spouse of the head. Taken together, a health shock to the head is far more influential than a health shock to the spouse.

Panels B of Tables 10-11 report the estimates on the heterogeneity effects by household and children characteristics. The estimates for each row come from a separate subsample. Therefore, these estimates are obtained from the most flexible model where the effects of all the covariates are allowed to be group specific (equivalent to including a full interaction between a subgroup

[34]In the expanded regression, I also control for interactions of each of the two health shocks with year and county fixed effects.

dummy and the covariates). When the sample is disaggregated for males and females separately, I find a stronger positive effect of health shocks on girls' employment than boys'. This suggests that labor supply among girls is more elastic to changes in household resources. Given that the schoolings of both boys and girls are negative affected by the health shock (Table 10), it appears that rural households increase the use of boys in household production and the use of girls in market-work. Providing health insurance has a greater impact on mitigating the negative impacts on school enrollment for girls than for boys: the estimated coefficient α_1 suggests that the reform has a stronger impact on maintaining school enrollment among girls than among boys, although the standard errors are large to reject the hypothesis that the male and female coefficients are identical.

In the remaining of Panels B, I compare the effects of the health shocks by age of the child (whether age 13 or above), education level of the head of the household, whether the child has any elder sibling (aged between 7 and 18 in the first year of the panel), household structure (whether there are more than two adult members in the household) and availability of a primary or middle school in the village. There are some interesting patterns contrasting the differences among groups, although some of the estimates are imprecise and not statistical different between groups because they come from a separate subsample. For instance, among children who do not have an older sibling or come from a household whose head is low-educated, health shock without health insurance has a large negative impact on schooling. Consequently, the role of health insurance provision in mitigating the negative effect of schooling is strong. The negative impact of a health shock on schooling is also significant and large among young children living in single-family household. This finding suggests that, relative to single-family households, multi-family households are better insured against health shock even without access to health insurance. On children's employment, it appears that health shock has the stronger impact among older children, children living in household headed by a high-educated adult or in villages without a primary or middle school. Consequently, the role of health insurance provision in mitigating the positive effect of health shock on employment is strong among these groups of children.

The evidence on human capital investment is in line with the predictions from models of child labor in which households face liquidity constraints. For instance, in Basu and Van (1998) and Baland and Robinson (2000), it can be optimal for households to have children working, even when keeping children in school leads to higher future wages. This is because keeping children out of the labor force (and in school) can incur a large utility cost from foregone household consumption.

Consistent with this prediction, I find that, without access to health insurance, households to increase the use of child employment and decrease child schooling following a health shock. The negative impact on schooling is larger among households which are more liquidity constrained (e.g. headed by low-educated adult or a single-family household), and access to health insurance is a particularly useful channel for these households to smooth against negative health shocks.

Existing evidence from other developing countries suggests that outcomes and well-beings of girls are more sensitive than boys to changes in household economic resources.[35] Consistent with this literature, this paper finds that there are gender-differences in the effect of health shock (and consequently the value of health insurance in mitigating the effect), particularly in terms of childrens employment. In China, the fraction of late-teens taking employment in cities is much higher among girls than boys, perhaps due to the skill content of the jobs for migrants (Meng, 2012). Therefore, girls employment is more elastic to health shocks given the high potential return from girls relative to boys.

A final important question is whether the estimated disruption in schooling from a household health shock has long-term impact on their level of education. In the data, very few children return to school after dropout. For these children in rural areas, even if the schooling disruption is transitory, this may lead to long-term impact on their level of education. Part of this is due to the "guest worker" system in China with controls over the type of jobs rural migrants are allowed and the social welfare and social services to which migrants are entitled (Meng 2012). In the data, over 40 percent of drop-out children are employed in cities and fall into the guest workers system which forbid them from complete education in urban areas. The rigid educational and labor market system means that a short-term disruption in schooling could have lasting impacts.

6.2 Investment in agricultural activities

How did the health shock affect household investment in agricultural activities? If households have complete access to credit, transitory shocks should have no effect on such investments, as borrowing allows households to separate the timing of investment from the timing of income. However, when household investments require fixed costs to be paid in advance of the investment returns and when households face credit constraints, the timing of household investments may depend on current income realizations (Yang, 2008).

[35]For instance, see Foster (1995) (well-being), Rose (1999) (child mortality), and Björkman-Nyqvist (2013) (education and school performance).

Table 12 presents the coefficient estimates on the health shock and the health shock interacted with reform, when the outcome variables are the change in household investment in farming, that in livestock, and the change in an indicator for whether the household had added any productive equipment since the last survey. Columns (1)–(2) present the estimates on farming and livestock investment. Focusing on the estimates from the health shock defined on the head and the spouse of the head, I find that prior to the reform, a 10-percentage-point increase in the size of the health shock is associated with a differential decrease in investment in farming of 6% and a differential decrease in investment in livestock of 12%. These estimates suggest that households adjust economic activities as a channel to insure against health shock. The fact that livestock investment is reduced the most indicates that households are shifting away from risky production (i.e., raising livestock) to conservative production in order to protect themselves from adverse income shocks (Rosenzweig and Wolpin, 1993). The estimated coefficients α_1 (on the interaction term between health shock and reform) are positive but insignificantly different from zero. Column (3) presents estimates when the outcome is the change in an indicator for positive livestock investment. There is little evidence of an adjustment in livestock at the extensive margin following a health shock. Changes in the indicator variable are used as outcome variables in the last column in Table 12. Columns (4) shows that, without access to health insurance, a 10-percentage-point increase in the size of the health shock is associated with a differential decrease in the likelihood of investing in productive equipment of 1 percentage point. The estimated coefficient on the reform indicator interacted with the health shock is positive and significantly different from zero. The predicted effect of a health shock after the reform is slightly positive, which indicates that with access to health insurance, the household no longer delays the purchase of productive equipment in response to a health shock.

7 Specification and Robustness Checks

Exogeneity of the health shock. One of the identifying assumptions is that the health shock is uncorrelated to the error term in equation (3), conditional on the included covariates in the model. As mentioned previously, health shocks may be correlated with unobserved changes to household resources, even with the first-differenced specification of the model. To assess whether correlated changes in household economic conditions are contributing to the estimated effect of the health shock, I examine the correlation between lagged income changes and levels, and the current health shock. If past economic shocks (such as job loss) affect current health, one would expect a negative

correlation between the health shock and lagged income changes. Alternatively, if past levels of income affect investment in health differently between the rich and the poor, one would also expect a negative correlation between the current health shock and the lagged income levels.

In Table 13, I show that there is little correlation between past income changes or levels and the current health shock, indicating that this potential source of bias may be less of a concern in the sample of study. For instance, in columns (1)–(2), the coefficients on lagged level (column 1) and lagged growth (column 2) of earnings of the head and spouse are small and insignificantly different from zero. The inclusion of county-by-year fixed effects potentially accounts for many confounding factors affecting the variations in health shocks.

Even if the effect of the health shock is contaminated by the effect of other types of shocks to household resources, the estimated coefficient on the interaction between the health shock and the reform (α_1) may be unbiased as long as the omitted variable bias in the coefficient of $\triangle h_{ijt}$ does not change with the reform. This is more of a concern if the reform changes the unobserved composition among households suffering from health shocks. Columns (3) reports the coefficient estimate on the reform indicator, controlling for county and year fixed effects, when the outcome variables is a dichotomous variable indicating whether the head or the spouse of the head reported any illness in the past four weeks (both severe and non-severe). The estimated coefficient on reform is small and insignificantly different from zero, suggesting no evidence that the reform changes the percentage share of households with illness. In column (4), among households reporting some illness, I regress the duration of illness on the reform indicator, controlling for the year and county fixed effects. I also find no evidence that the reform shifts the duration of illness. Columns (7)–(8) confirm this finding by repeating the analysis on the health shock measured defined on the head, the spouse of the head and other working-age household members. Overall, these results indicate that the reform is unlikely to affect the composition of households with illness.

Alternative measures of health shock. The measure of health shock considered so far has a recall period of four weeks ("number of days with severe sickness in the past four weeks"). Income and agricultural investments have a long recall period (with respect to the past 12 months). The difference in recall period may imply that the outcome variable may appear "excessively smooth" to the health shock (coefficient α_0 biased toward zero), especially if the impact of the health shock in the past four weeks is small relative to the level of outcome prior to the health shock. In Appendix Table A2, I report estimates of regression 3 using changes in the weeks of sickness affecting work

and daily activities over the past three months as a measure of health shock. For better comparison the marginal effect of a health shock to the baseline specification, the number of weeks of illness is normalized as the fraction of weeks with sickness in the past three months (assuming 12 weeks). In general, across most outcomes of interest, the estimated effects of a negative health shock are quite similar to the estimates from the baseline model. The estimated coefficients on the interaction term between the health shock and the reform are qualitatively similar but imprecisely estimated. This alternative measure of health shock is only available until 2006.[36] This implies that a reduction in sample size (by over 30%) and, given that less than 60 percent of households were under the health insurance reform by 2006, a reduction in the effective sample size needed to identify the effect of the reform interacted with health shocks. Furthermore, if the effect of the reform is stronger over time (e.g. as people learn about the program and the program becomes more generous over time), the fact that we do not observe outcome variables beyond 2006 may lead to under-estimate the effect of the program. All these factors contribute to large standard errors on the effects of the reform on household choices against a health shock.[37]

8 Conclusion

A negative health shock could incur two types of costs on an individual. One is the direct negative effect on health and productivity, which in turn, affect individual earnings. The other type of cost is the out-of-pocket medical expenditure associated with treating the illness. Without access to health insurance, family labor supply (including child labor), transfer and subsidies, nature of the economic activities may be important mechanisms to smooth total household income against individual health shocks. Household may also resort to other channels of self-insurance in order to smooth household consumption. Health insurance may lessen the shock to individual earnings through high quality treatment of the illness, reduce out-of-pocket medical expenditure, and interact with existing self-insurance mechanisms against health shocks.

[36] Starting from the 2009 survey, these questions are no longer included in the survey.

[37] Another potential measure of health shock is change in any self-reported chronic disease. Starting from 1997, information on the following types of doctor diagnosis of chronic diseases is available in each survey: high blood pressure, diabetes, myocardial infarction, apoplexy, and bone fracture. Unlike the health shocks defined previously, the CNHS does not ask the respondent the severity of these chronic diseases. Therefore, the corresponding health shock defined by chronic disease may not be severe enough to have significant economic impact on work capacity and household resources. For instance, a substantial fraction of individuals reporting chronic disease does not have any difficulty in carrying out work and daily activities. Another poential problem with the chronic disease measure is that the survey only asked whether the respondent has any doctor diagnosis of chronic disease (the questions begin with "Has a doctor ever told you that..."). Given that the reform had an impact on regular health check-up as documented in the literature (e.g. Lei and Lin (2009)), health shock defined by chronic diseases is potentially endogenous to the reform.

Using the rollout of universal health insurance in rural China as a source of exogenous variation in coverage, this paper provided empirical evidence on the role of public health insurance in mitigating adverse outcomes associated with health shocks. Even in the absence of health insurance coverage, a health shock was found to affect neither growth in household income and nor growth in household consumption. In the meanwhile, a health shock had sizable impacts on individual earnings growth. For instance, a 10-percentage-points increase in the size of the health shock was associated with a differential decrease in total income of the head and spouse of 17 percent. These results supported the hypothesis that family labor supply is important channel households use to insure against health shocks. I found that the health insurance program crowds out important household choices such as children's schooling and child employment which were used to smooth consumption in the absence of health insurance. For instance, without access to public health insurance, 10-percentage-point increase in the size of the health shock was associated with a differential increase in child employment of 8 percentage points. After the reform, the net effect of a health shock with health insurance on school enrollment was small and statistically insignificant.

The results of this paper suggest that analyzing consumption fluctuations alone may provide an incomplete picture of the true value of additional social insurance. To understand the net welfare gain of public insurance programs, an important policy question is at what cost public insurance replaces private arrangements of smoothing against shocks. The existing literature has argued that additional public insurance programs may have little net effects on the welfare of the intended beneficiaries if they are already completely insured against shocks, because the provision of these programs may only crowd out any private arrangements for self-insurance. The evidence from the health insurance reform in rural China suggests that additional public insurance programs could lead to net gains by reducing the use of costly self-insurance mechanisms.

APPENDIX

A An Analytical Framework

I follow Chetty and Looney (2006) to outline a simple model highlighting the welfare gains from publicly provided insurance. Suppose that there are two states in the world, one with good health and the other with bad health. Suppose that, in the good state, the utility cost of obtaining consumption level c is $\theta_g c$. In the bad state, the utility cost of reaching consumption level c requires a larger cost $\theta_b c$. I normalize $\theta_g = 1$, so θ_b measures the additional utility cost of reaching consumption level c in the bad state, relative to the cost in the good state. If θ_b is high, households that are affected by shocks will need to sacrifice more in order to reach the same consumption level as before. To facilitate later discussion, suppose that the household can choose from a menu of insurance channels denoted by $\{\theta_b^1, \theta_b^2, \ldots\}$. θ_b^k summarizes the cost of consumption smoothing from different channels. A utility-maximizing household would choose $\theta_b = \min\{\theta_b^1, \theta_b^2, \ldots\}$. Assuming a constant relative risk aversion (CRRA) utility function, the optimal changes in consumption in response to a bad health shock can be written as

$$\frac{\triangle c_i}{c_i} = 1 - (\frac{1}{\theta_b})^{1/\gamma} \tag{1}$$

where γ is the coefficient of risk aversion. From equation (1), it is obvious that changes in consumption depend on two parameters: the cost of consumption smoothing and the coefficient of risk aversion. Consumption may then not fluctuate much in response to health shocks because of either a low cost of smoothing (such as easy access to credit markets) or a high cost of smoothing but with households that are very risk averse (γ is large).

B Variable Construction and Definitions

Health shock. Defined in Section 4 of the main text.

Household income and its components. Household income is the sum of nine potential sources of income in the last year: business, farming (including gardening), fishing, livestock, non-retirement wages, retirement income, subsidies, and other income (mainly transfers). Transfers include total private transfers received from parents, children, relatives, and friends in the past year (both cash transfers and self-reported value of in-kind transfers). Subsidies include the value of in-kind benefits, mostly from the employer. Individual income is the sum of seven potential sources of income: business, farming, fishing, livestock, non-retirement wages, and retirement income. Individual income from potential sources of business, farming, fishing, and livestock is calculated based on the individual proportion of net household income from each of these sources. The individual proportion is the time that each household member spends on each activity divided by the sum of the time that all household members spend on the activity. Therefore, for rural

household members, holding total household income fixed, the change in individual-level income from these sources mostly reflects the change in labor supply.

Food consumption. Food consumption is collected by using a dietary questionnaire recording items and the amounts of food consumption for each member of the household over the past three days (including food consumed both at and away from home). The CHNS provides calculated macronutrients based on the individual food consumption data. The macronutrients include the average daily intake of calories (in kilocalories), proteins (in grams), fats (in grams) and carbohydrates (in grams).

Out-of-pocket medical expenditure. Medical expenditure includes the cost of treating the illness (the same illness used to define the health shock) at up to two clinics or hospitals, as well as the cost of informal treatment if the individual chose to use informal care (i.e. did not go to a clinic). Out-of-pocket medical expenditure is the total medical expenditure multiplied by the share of expenses not covered by health insurance. Note that, when insurance covers all expenses, the CHNS did not record the cost of treatment. In this case, the out-of-pocket expenditure is zero.

Children's school enrollment. School enrollment is whether a child aged between 7 and 18 years attending school at the time of the survey. School enrollment status for each child is defined using a combination of two variables. The first and main variable is whether the child is currently enrolled in school. To track children who are members of the household but did not reside with the household at the time of the survey (e.g., temporary migrant workers and students at boarding schools), I rely on additional information provided by the household roster. On the household roster, for each member away from the household, CHNS records the reasons for not currently residing in the household. A child is defined as enrolled in school if the response to this question is "gone to school". A child is defined as not in school if the selected response is either "seeking employment in cities" or "in the military".

Child employment. I define child employment as whether a child (aged between 7 and 18 years) is working at the time of the survey. It is defined using a combination of two variables. One is whether an individual is reported as working (excluding household work) at the time of the survey. Prior to 2004, this variable was collected for household members aged 16 years and above. Starting from the 2004 wave, this variable is also collected for all children under the age of 16 years. In addition, I also use the question on the reasons given for not residing in the household to trace children employed away from home. Specifically, a child is defined as working if the response is "employment in cities".

Agricultural investments. Investment in farming refers to the total amount spent, in the past year, on leasing land, purchasing seedlings, fertilizer, tools, insecticides, and hiring labor. This variable is only available for households with farming and/or gardening activities on a collective,

state, or household farm. Investment in livestock includes the total amount spent on purchasing, feeding, and caring for up to four types of livestock and poultry. This variable is only available for households with activities associated with raising livestock or poultry. Productive equipments used in agriculture activities include tractors (of all sizes) and power thresher. The indicator for new productive equipment is equal to one if a household had added any of this equipment since the last survey.[38]

[38]Each wave surveys the quantity of each type of equipment owned by the household. Investment is based on whether there has been an increase in the aggregate quantity between the current wave and the previous wave.

References

Angelucci, M., and O. Attanasio (2013): "The Demand for Food of Poor Urban Mexican Households: Understanding Policy Impacts Using Structural Models," *American Economic Journal: Economic Policy*, 5(1), 146–205.

Angelucci, M., G. De Giorgi, and I. Rasul (2012): "Resource Pooling Within Family Networks: Insurance and Investment," *UCL Working paper*.

Attanasio, O., and J.-V. Rios-Rull (2000): "Consumption smoothing in island economies: Can public insurance reduce welfare?," *European Economic Review*, 44(7), 1225–1258.

Bai, C.-E., and B. Wu (2014): "Health insurance and consumption: evidence from Chinas new cooperative medical scheme," *Journal of Comparative Economics*, 42(2), 450–469.

Baicker, K., A. Finkelstein, J. Song, and S. Taubman (2013): "The Impact of Medicaid on Labor Force Activity and Program Participation: Evidence from the Oregon Health Insurance Experiment," Working Paper 19547, National Bureau of Economic Research.

Baland, J.-M., and J. A. Robinson (2000): "Is child labor inefficient?," *Journal of Political Economy*, 108(4), 663–679.

Basu, K., and P. H. Van (1998): "The economics of child labor," *American economic review*, pp. 412–427.

Beegle, K., R. H. Dehejia, and R. Gatti (2006): "Child labor and agricultural shocks," *Journal of Development Economics*, 81(1), 80–96.

Björkman-Nyqvist, M. (2013): "Income shocks and gender gaps in education: Evidence from Uganda," *Journal of Development Economics*, 105, 237–253.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2012): "Losing heart? The effect of job displacement on health," Discussion paper, National Bureau of Economic Research.

Brown, P. H., and A. Park (2002): "Education and poverty in rural China," *Economics of Education Review*, 21(6), 523–541.

Chen, Y., and G. Z. Jin (2012): "Does health insurance coverage lead to better health and educational outcomes? Evidence from rural China," *Journal of health economics*, 31(1), 1–14.

Chetty, R. (2006): "A general formula for the optimal level of social insurance," *Journal of Public Economics*, 90(10), 1879–1901.

Chetty, R., and A. Looney (2006): "Consumption smoothing and the welfare consequences of social insurance in developing economies," *Journal of Public Economics*, 90(12), 2351–2356.

——— (2007): "Income risk and the benefits of social insurance: Evidence from Indonesia and the United States," in *Fiscal Policy and Management in East Asia, NBER-EASE, Volume 16*, pp. 99–121. University of Chicago Press.

Chou, S.-Y., J.-T. Liu, and J. K. Hammitt (2003): "National health insurance and precautionary saving: evidence from Taiwan," *Journal of Public Economics*, 87(9), 1873–1894.

Cochrane, J. H. (1991): "A Simple Test of Consumption Insurance," *The Journal of Political Economy*, 99(5), 957–976.

Cox, D., Z. Eser, and E. Jimenez (1998): "Motives for private transfers over the life cycle: An analytical framework and evidence for Peru," *Journal of Development Economics*, 55(1), 57–80.

Currie, J., and B. C. Madrian (1999): "Health, health insurance and the labor market," *Handbook of labor economics*, 3, 3309–3416.

Cutler, D. M., and J. Gruber (1996): "Does public insurance crowd out private insurance?," *The Quarterly Journal of Economics*, 111(2), 391–430.

Department of Health, C. (2006): "On Expanding the New Cooperative Medical Scheme (in Chinese)," *Bulletin of the Department of Health of China*, 2006(13).

——— (2008): "On Strengthening the New Cooperative Medical Scheme in 2008 (in Chinese)," *Bulletin of the Department of Health of China*, 2008(17).

Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J. P. Newhouse, H. Allen, K. Baicker, et al. (2012): "The Oregon Health Insurance Experiment: Evidence from the First Year*," *The Quarterly Journal of Economics*, 127(3), 1057–1106.

Foster, A. D. (1995): "Prices, credit markets and child growth in low-income rural areas," *The Economic Journal*, pp. 551–570.

Gertler, P., and J. Gruber (2002): "Insuring Consumption Against Illness," *American economic review*, 92(1), 51–70.

Gertler, P., D. I. Levine, and E. Moretti (2009): "Do microfinance programs help families insure consumption against illness?," *Health economics*, 18(3), 257–273.

Gruber, J., and A. Yelowitz (1999): "Public health insurance and private savings," *Journal of Political Economy*, 107(6), 1249.

Islam, A., and P. Maitra (2012): "Health shocks and consumption smoothing in rural households: Does microcredit have a role to play?," *Journal of development economics*, 97(2), 232–243.

Jalan, J., and M. Ravallion (1999): "Are the poor less well insured? Evidence on vulnerability to income risk in rural China," *Journal of development economics*, 58(1), 61–81.

Jensen, R. T. (2004): "Do private transfers displace the benefits of public transfers? Evidence from South Africa," *Journal of Public Economics*, 88(1), 89–112.

Kochar, A. (1995): "Explaining household vulnerability to idiosyncratic income shocks," *The American Economic Review*, 85(2), 159–164.

Lei, X., and W. Lin (2009): "The new cooperative medical scheme in rural China: Does more coverage mean more service and better health?," *Health Economics*, 18(S2), S25–S46.

Mace, B. J. (1991): "Full Insurance in the Presence of Aggregate Uncertainty," *Journal of Political Economy*, 99(5), 928–56.

Meng, X. (2012): "Labor Market Outcomes and Reforms in China," *The Journal of Economic Perspectives*, 26(4), 75–101.

Mohanan, M. (2013): "Causal effects of health shocks on consumption and debt: quasi-experimental evidence from bus accident injuries," *Review of Economics and Statistics*, 95(2), 673–681.

Morduch, J. (1995): "Income Smoothing and Consumption Smoothing," *The Journal of Economic Perspectives*, 9(3), 103–114.

Rose, E. (1999): "Consumption smoothing and excess female mortality in rural India," *Review of Economics and statistics*, 81(1), 41–49.

Rosenzweig, M. R., and K. Wolpin (1993): "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investment in Bullocks in India," *Journal of Political Economy*, 101(2), 223–44.

State Council, C. (2003): "Guidlines Establishing the New Cooperative Medical Scheme (in Chinese)," *Bulletin of the State Council of China*.

Strauss, J., P. Gertler, O. Rahman, and K. Fox (1993): "Gender and Life-Cycle Differentials in the Patterns and Determinants of Adult Health," *Journal of Human Resources*, 28(4), 791–837.

Townsend, R. M. (1994): "Risk and Insurance in Village India," *Econometrica*, 62(3), 539–591.

Wagstaff, A., M. Lindelow, G. Jun, X. Ling, and Q. Juncheng (2009): "Extending health insurance to the rural population: An impact evaluation of China's new cooperative medical scheme," *Journal of health economics*, 28(1), 1–19.

Wagstaff, A., and M. Pradhan (2005): "Health Insurance Impacts on Health and Nonmedical Consumption in a Developing Country," *World Bank Policy Research Working Paper*, (3563).

Yang, D. (2008): "International migration, remittances and household investment: Evidence from philippine migrants exchange rate shocks*," *The Economic Journal*, 118(528), 591–630.

Yip, W., and W. C. Hsiao (2008): "The Chinese health system at a crossroads," *Health Affairs*, 27(2), 460–468.

Yu, X., and D. Abler (2009): "The demand for food quality in rural China," *American Journal of Agricultural Economics*, 91(1), 57–69.

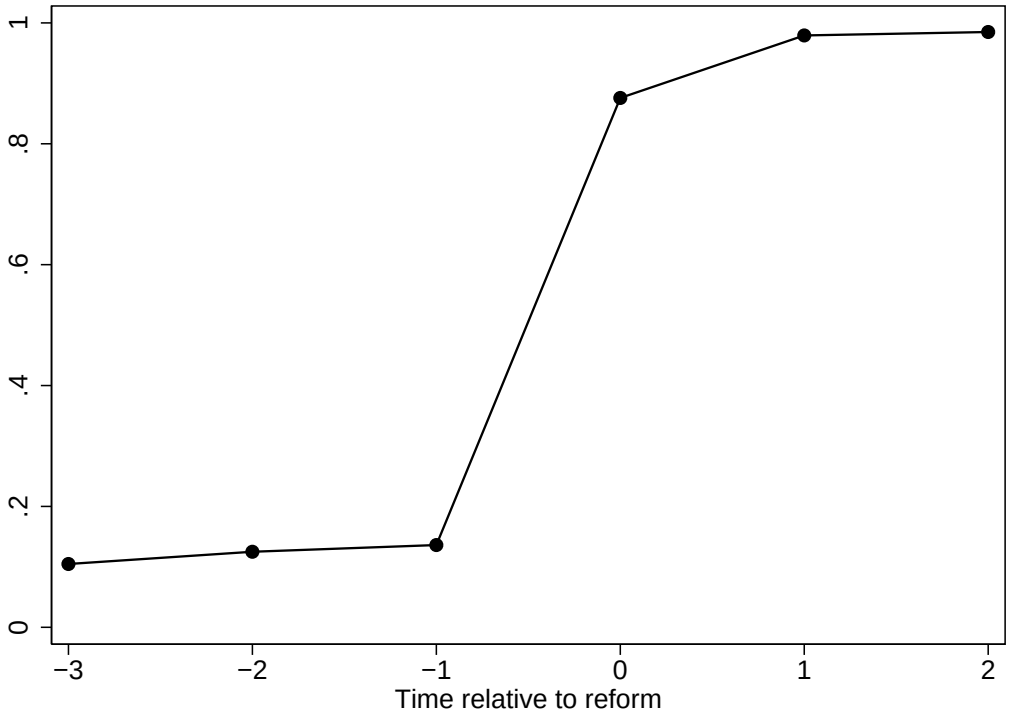

Figure 1: Expansion of Health Insurance Coverage

Notes: For each county, I recenter the data such that time zero is equal to the first wave observed since the reform. The graph plots average rate of health insurance coverage (y-axis) on time (in survey waves) relative to the year of health insurance reform (x-axis). The y-axis is the percentage of household covered by any health insurance plan.

Table 1: Implementation of the Reform By County

	Counties		
Year	Total	Treated	% treated
1997	32	0	0.0%
2000	36	0	0.0%
2004	36	3	8.3%
2006	36	22	61.1%
2009	36	36	100.0%
2011	36	36	100.0%

Note: In the sample, each round of the survey contains households sampled from 36 counties across 9 provinces (four counties per province). Four counties from Liaoning province were not surveyed in the 1997 wave. Community-level survey data from the CHNS are used to determine the year of implementation of the NCMS at county level.

Table 2: Implementation of the Reform and Household Enrollment

				Households			
Year	Total	Access to NCMS	% access	Insured by CMS[a]	% insured by CMS[a]	Insured	% insured
1997	1421	0	0.0%	182	12.8%	210	14.8%
2000	1567	0	0.0%	110	7.0%	152	9.7%
2004	1545	119	7.7%	225	14.6%	265	17.2%
2006	1525	907	59.5%	785	51.5%	814	53.4%
2009	1508	1508	100.0%	1410	93.5%	1439	95.4%
2011	1451	1451	100.0%	1373	94.6%	1407	97.0%

Note: a. CMS includes both the CMS existed prior to the reform and the NCMS program introduced by the reform. Prior to the 2009 round, the micro-level data from CHNS did not make a distinction between the old CMS and the NCMS coverage.

Table 3: Characteristics of Counties and the Timing of Reform

	Year of reform
Number of days with severe illness	0.046
	(0.182)
Log(average household income)	1.164
	(1.066)
Growth in household income	-66.688
	(46.982)
Share of children in school	-2.434
	(4.271)
Growth in the share of children in school	34.818
	(159.019)
Any health insurance	-1.671
	(1.590)
Head with at least 9 years of schooling	-2.977
	(2.931)
Age of household head	0.038
	(0.131)
Constant	1998.6
	(10.874)
R^2	0.219
N	36

Note: This table presents the coefficient estimates in an OLS regression at the county level. The dependent variable is the calendar year of the reform in a given county. The covariates are the average values at the county level in year 2000 (the last pre-reform year among all counties), including number of days of severe sickness, log household income, growth in household income, share of children in school, growth in the share of children in school, proportion of households with any health insurance, and level of education and age of the household head.

Table 4: Household Characteristics

	Mean	SD	Observations
Age of the head	47.01	9.98	10198
Head being single	0.06	0.24	10198
Head with at least 9 years of education	0.66	0.47	10198
Household size	3.88	1.44	10198
Total household income	17280.41	26510.24	10198
Household income exclu. transfers and subsidies	15455.57	25393.53	10198
Head and spouse joint earnings	11616.53	21278.28	10198
Household income from farming	6351.23	9013.43	10198
Household income from livestock	597.08	4350.88	10198
Investment in farming[a]	1272.18	1661.57	8312
Investment in livestock[b]	745.86	1576.97	5472
Added productive equipment[c]	0.10	0.30	9051
Share of children (aged 7-18) in school[d]	0.78	0.38	4961
Share of children at work (aged 7-18)[d]	0.16	0.34	4961
Calories (in kilocalorie)	2351.64	628.35	10109
Protein (in grams)	67.94	21.00	10094
Fat (in grams)	65.70	32.40	10094
Carbohydrate (in grams)	362.20	114.87	10096
Share of days with severe illness (head and spouse)	2.27	11.94	10198
Share of days with severe illness (head and spouse and working members)	2.45	12.54	10198

Notes:

a. For households with farming activities on a collective, state or a household farm.

b. For households with activities in raising livestock or poultry.

c. For households who have reported the quantity of productive equipment in the previous wave.

d. For households with children aged between 7 and 18.

Table 5: Impact of Health Shocks and Health Insurance on Income

Outcomes: change in ...	Regression Estimates α_0	α_1	$H_0 : \alpha_0 + \alpha_1 = 0$ (p-value)	Number of observations
(a) Health shock defined on head and spouse				
Log (head and spouse income)	-0.017***	0.009	0.011	6432
	(0.002)	(0.005)		
Log (household income)	0.001	-0.002	0.824	7366
	(0.002)	(0.005)		
Log (household income, excl.	-0.002	0.003	0.755	7104
subsidies and transfer income)	(0.002)	(0.003)		
Log (income from	0.026**	-0.034	0.615	2320
other members)	(0.011)	(0.028)		
(b) Health shock defined on head, spouse and working age members				
Log (income from head, spouse	-0.004*	0.003	0.806	6608
and working age members)	(0.002)	(0.005)		
Log (household income)	0.001	-0.003	0.404	7366
	(0.002)	(0.005)		
Log (household income, excl.	-0.001	0.000	0.629	7104
subsidies and transfer income)	(0.002)	(0.003)		

*** p<0.01, ** p<0.05, * p<0.1. Each cell in regression columns 1 and 2 presents the coefficient estimates on health shock (α_0) and health shock interacted with the reform (α_1) in an OLS regression as specified in equation (3). Column 3 presents p-values of testing the null hypothesis that $\alpha_0 + \alpha_1 = 0$. Standard errors (in parentheses) are clustered at the county level. All dependent variables are first-differenced variables. In Panel (a), health shock is the change in the fraction of the past 28 days (in percentage points) in which the head and the spouse of head suffered from severe illness limiting daily activities. In Panel (b), the health shock is the change in the fraction of the past 28 days (in percentage points) in which the head, the spouse of head and other working-age household members (aged 18-55) suffered from severe illness limiting daily activities. Each regression includes the same set of covariates discussed in Section 3. Details of the variable construction and definitions are described in Appendix B.

Table 6: Impact of Health Shocks and Health Insurance on Consumption Growth

Outcomes: change in ...	Regression Estimates α_0	α_1	$H_0 : \alpha_0 + \alpha_1 = 0$ (p-value)	Number of observations
(a) Health shock defined on head and spouse				
Log (caloric intake)	-0.000	-0.000	0.284	7491
	(0.000)	(0.001)		
Log (protein intake)	-0.000	0.000	0.991	7472
	(0.001)	(0.001)		
Log (fat intake)	-0.001	0.000	0.721	7473
	(0.001)	(0.002)		
Log (carbohydrate intake)	-0.000	-0.001	0.027	7474
	(0.000)	(0.001)		
(b) Health shock defined on head, spouse and working age members				
Log (caloric intake)	-0.000	0.000	0.851	7491
	(0.001)	(0.001)		
Log (protein intake)	-0.000	0.001	0.757	7472
	(0.001)	(0.001)		
Log (fat intake)	-0.001	0.001	0.769	7473
	(0.001)	(0.003)		
Log (carbohydrate intake)	-0.000	-0.001	0.122	7474
	(0.000)	(0.001)		

*** p<0.01, ** p<0.05, * p<0.1. Each cell in regression columns 1 and 2 presents the coefficient estimates on health shock (α_0) and health shock interacted with the reform (α_1) in an OLS regression as specified in equation (3). Column 3 presents p-values of testing the null hypothesis that $\alpha_0 + \alpha_1 = 0$. Standard errors (in parentheses) are clustered at the county level. All dependent variables are first-differenced variables. Household consumption is defined as the average macronutrient intake among adult members (aged 18+) of the household. In Panel (a), the health shock is the change in the fraction of the past 28 days (in percentage points) in which the head and the spouse of head suffered from severe illness limiting daily activities. In Panel (b), the health shock is change in the fraction of the past 28 days (in percentage points) in which the head, the spouse of head and other working-age household members (aged 18-55) suffered from severe illness limiting daily activities. Each regression includes the same set of covariates discussed in Section 3. Details of the variable construction and definitions are described in Appendix B.

Table 7: Impact of Health Shocks and Health Insurance on Consumption Growth: Excluding the Consumption of Head/Spouse with Illness

Outcomes: change in ...	Regression Estimates α_0	α_1	$H_0 : \alpha_0 + \alpha_1 = 0$ (p-value)	Number of observations
Health shock defined on head and spouse				
Log (caloric intake)	-0.001	-0.000	0.486	6762
	(0.001)	(0.002)		
Log (protein intake)	0.001	-0.001	0.487	6742
	(0.001)	(0.002)		
Log (fat intake)	-0.004**	0.006*	0.239	6744
	(0.001)	(0.003)		
Log (carbohydrate intake)	-0.000	-0.002	0.077	6739
	(0.001)	(0.002)		

*** p<0.01, ** p<0.05, * p<0.1. Each cell in regression columns 1 and 2 presents the coefficient estimates on health shock (α_0) and health shock interacted with the reform (α_1) in an OLS regression as specified in equation (3). Column 3 presents p-values of testing the null hypothesis that $\alpha_0 + \alpha_1 = 0$. Standard errors (in parentheses) are clustered at the county level. All dependent variables are first-differenced variables. Household consumption is defined as the average macronutrient intake among adult members (aged 18+) of the household, but excluding the consumption from the head and/or spouse who suffered from any severe illness. Health shock is the change in the fraction of the past 28 days (in percentage points) in which the head and the spouse of the head suffered from severe illness limiting daily activities. Each regression includes the same set of covariates discussed in Section 3. Details of the variable construction and definitions are described in Appendix B.

Table 8: Effects of the Reform on Health Expenditures

	Positive out-of-pocket expenditure	Log out-of-pocket expenditure	Out-of-pocket expenditure (Tobit)
	(1)	(2)	(3)
(a) Health shock defined on head and spouse			
Reform	-0.049***	0.278	-569.835***
	(0.017)	(0.397)	(57.449)
N	10082	1199	10082
(b) Health shock defined on head, spouse and working age members			
Reform	-0.052***	0.224	-504.626***
	(0.018)	(0.394)	(51.549)
N	10127	1288	10127

*** p<0.01, ** p<0.05, * p<0.1. Within each panel, each cell in the first row presents the coefficient estimates on the Reform indicator in a regression as specified in equation (6). Coefficients in the first two columns are estimated by OLS. Coefficients in the last column is estimated by Tobit. Standard errors (in parentheses) are clustered at the county level. In Panel (a), health shock is the change in the fraction of the past 28 days (in percentage points) in which the head and the spouse of head suffered from severe illness limiting daily activities. In Panel (b), health shock is the change in the fraction of the past 28 days (in percentage points) in which the head, the spouse of the head and other working-age household members (aged 18-55) suffered from severe illness limiting daily activities. Each regression includes the same set of covariates discussed in Section 5.2. Details of the variable construction and definitions are described in Appendix B.

Table 9: Impact of Health Shocks and Health Insurance on Human Capital Investment

Outcomes: change in ...	Schooling	Employment
(a) Health shock defined on head and spouse		
α_0	-0.006***	0.008***
	(0.001)	(0.001)
α_1	0.005**	-0.006**
	(0.002)	(0.002)
$H_0 : \alpha_0 + \alpha_1 = 0$		
(p-value)	0.852	0.128
(b) Health shock defined on head, spouse and working age members		
α_0	-0.005***	0.008***
	(0.001)	(0.001)
α_1	0.005**	-0.006**
	(0.002)	(0.002)
$H_0 : \alpha_0 + \alpha_1 = 0$		
(p-value)	0.666	0.342
N	3974	4308

*** p<0.01, ** p<0.05, * p<0.1. Within each panel, each cell in rows 1 and 2 presents the coefficient estimates on health shock (α_0) and health shock interacted with the reform (α_1) in an OLS regression as specified in equation (3). Row 3 presents p-values of testing the null hypothesis that $\alpha_0 + \alpha_1 = 0$. Standard errors (in parentheses) are clustered at the county level. All dependent variables are first-differenced variables. In Panel (a), health shock is the change in the fraction of the past 28 days (in percentage points) in which the head and the spouse of the head suffered from severe illness limiting daily activities. In Panel (b), health shock is the change in the fraction of the past 28 days (in percentage points) in which the head, the spouse of the head and other working-age household members (aged 18-55) suffered from severe illness limiting daily activities. Each regression includes the same set of covariates discussed in Section 3. Details of the variable construction and definitions are described in Appendix B.

Table 10: Heterogeneous Impact of Health Shocks and Health Insurance: Schooling of Children

Outcomes: change in ...	Regression Estimates α_0	α_1	$H_0 : \alpha_0 + \alpha_1 = 0$ (p-value)	Number of observations
A. Single Regression				
Health shock to the head	-0.004*	0.005	0.557	3638
	(0.002)	(0.003)		
Health shock to the spouse of the head	-0.001	0.001	0.806	3638
	(0.001)	(0.002)		
B. Separate Regressions				
1. Girls	-0.006*	0.012	0.166	1823
	(0.003)	(0.007)		
Boys	-0.006***	0.000	0.001	2151
	(0.001)	(0.003)		
2. Young	-0.011***	0.002	0.002	1035
	(0.002)	(0.001)		
Old	-0.006***	0.005	0.449	2939
	(0.002)	(0.004)		
3. Low-education mother	-0.019**	0.018	0.886	951
	(0.007)	(0.011)		
High-education mother	-0.002**	0.004	0.289	3023
	(0.001)	(0.003)		
4. Has elder sibling	-0.005**	0.004	0.874	2067
	(0.002)	(0.004)		
No elder sibling	-0.010***	0.010***	0.913	1907
	(0.001)	(0.003)		
5. Multi-family home	0.004	0.004	0.016	1473
	(0.004)	(0.004)		
Single-family home	-0.014***	0.011**	0.455	2501
	(0.002)	(0.005)		
6. School in the village	-0.006***	0.005*	0.415	3040
	(0.002)	(0.003)		
No school in the village	-0.010	-0.001	0.192	934
	(0.017)	(0.021)		

*** p<0.01, ** p<0.05, * p<0.1. In Panel A, each cell in regression columns 1 and 2 presents the coefficient estimates on the corresponding health shock (shock to the head and the spouse of the head) and the corresponding health shock interacted with the reform (α_1) from a single regression. In Panel B, each cell in regression columns 1 and 2 presents the coefficient estimates on health shock (α_0) and health shock interacted with the reform (α_1) in an OLS regression as specified in equation (3) on a different subsample. Standard errors (in parentheses) are clustered at the county level. All dependent variables are first-differenced variables. Health shock is the change in the fraction of the past 28 days (in percentage points) in which the head and the spouse of the head suffered from severe illness limiting daily activities. Each regression includes the same set of covariates discussed in Section 3. Details of the variable construction and definitions are described in Appendix B.

Table 11: Heterogeneous Impact of Health Shocks and Health Insurance: Employment of Children

Outcomes: change in ...	Regression Estimates α_0	α_1	$H_0 : \alpha_0 + \alpha_1 = 0$ (p-value)	Number of observations
A. Single Regression				
Health shock to the head	0.005***	-0.006*	0.465	3947
	(0.002)	(0.004)		
Health shock to the spouse of the head	0.001	-0.003	0.269	3947
	(0.001)	(0.002)		
B. Separate Regressions				
1. Girls	0.007**	-0.011*	0.313	1989
	(0.003)	(0.006)		
Boys	0.001	-0.001	0.668	2319
	(0.001)	(0.002)		
2. Young	0.001	-0.000	0.396	1148
	(0.001)	(0.000)		
Old	0.006***	-0.008***	0.051	3160
	(0.002)	(0.003)		
3. Low-education mother	0.001	-0.008	0.170	1034
	(0.006)	(0.010)		
High-education mother	0.007***	-0.006**	0.449	3274
	(0.001)	(0.002)		
4. Has elder sibling	0.004**	-0.007**	0.035	2235
	(0.002)	(0.003)		
No elder sibling	0.001	-0.007**	0.007	2073
	(0.002)	(0.003)		
5. Multi-family home	0.003	-0.009*	0.106	1587
	(0.004)	(0.005)		
Single-family home	0.004**	-0.010*	0.113	2721
	(0.002)	(0.005)		
6. School in the village	0.002**	-0.007**	0.018	3320
	(0.001)	(0.003)		
No school in the village	0.023***	-0.028***	0.022	988
	(0.005)	(0.005)		

*** p<0.01, ** p<0.05, * p<0.1. In Panel A, each cell in regression columns 1 and 2 presents the coefficient estimates on the corresponding health shock (shock to the head and the spouse of the head) and the corresponding health shock interacted with the reform (α_1) from a single regression. In Panel B, each cell in regression columns 1 and 2 presents the coefficient estimates on health shock (α_0) and health shock interacted with the reform (α_1) in an OLS regression as specified in equation (3) on a different subsample. Standard errors (in parentheses) are clustered at the county level. All dependent variables are first-differenced variables. Health shock is the change in the fraction of the past 28 days (in percentage points) in which the head and the spouse of the head suffered from severe illness limiting daily activities. Each regression includes the same set of covariates discussed in Section 3. Details of the variable construction and definitions are described in Appendix B.

Table 12: Impact of Health Shocks and Health Insurance on Investment in Agricultural Activities

Outcomes: change in ...	Log (invest. in farming)	Log (invest. in livestock)	Indicator for livestock invest.	Indicator for new prod. equip.
(a) Health shock defined on head and spouse				
α_0	-0.006**	-0.012***	0.001	-0.001*
	(0.002)	(0.003)	(0.001)	(0.000)
α_1	0.001	0.007	0.000	0.002**
	(0.005)	(0.007)	(0.002)	(0.001)
Test of $\alpha_0 + \alpha_1 = 0$				
(p-value)	0.124	0.163	0.255	0.016
(b) Health shock defined on head, spouse and working age members				
α_0	-0.003	-0.010**	0.000	-0.001**
	(0.002)	(0.004)	(0.001)	(0.000)
α_1	-0.001	0.002	0.001	0.002**
	(0.005)	(0.008)	(0.002)	(0.001)
Test of $\alpha_0 + \alpha_1 = 0$				
(p-value)	0.115	0.087	0.172	0.015
N	5565	2471	7609	7429

*** p<0.01, ** p<0.05, * p<0.1. Within each panel, each cell in rows 1 and 2 presents the coefficient estimates on health shock (α_0) and health shock interacted with the reform (α_1) in an OLS regression as specified in equation (3). Row 3 presents p-values of testing the null hypothesis that $\alpha_0 + \alpha_1 = 0$. Standard errors (in parentheses) are clustered at the county level. All dependent variables are first-differenced variables. In Panel (a), health shock is the change in the fraction of the past 28 days (in percentage points) in which the head and the spouse of the head suffered from severe illness limiting daily activities. In Panel (b), health shock is the change in the fraction of the past 28 days (in percentage points) in which the head, the spouse of the head and other working-age household members (aged 18-55) suffered from severe illness limiting daily activities. Each regression includes the same set of covariates discussed in Section 3. Details of the variable construction and definitions are described in Appendix B.

Table 13: Specification Checks: Past Income Changes, Current Health Shock and the Reform

	Health of head and spouse				Health of head, spouse and working age members			
	$\triangle h_{ijt}$	$\triangle h_{ijt}$	Any illness	Duration	$\triangle h_{ijt}$	$\triangle h_{ijt}$	Any illness	Duration
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lag level of earnings	0.028				0.053			
	(0.223)				(0.223)			
Lag growth of earnings		0.026				-0.126		
		(0.166)				(0.163)		
Reform			-0.011	0.441			-0.014	-0.926
			(0.026)	(3.398)			(0.027)	(3.656)
N	6995	4848	10082	1717	6995	4848	10127	1818

*** p<0.01, ** p<0.05, * p<0.1. Each column presents the coefficient estimates in a separate OLS regression. Standard errors (in parentheses) are clustered at the county level. In columns 1–4, health shock is the change in the fraction of the past 28 days (in percentage points) in which the head and the spouse of the head suffered from severe illness limiting daily activities. In columns 5–8, health shock is the change in the fraction of the past 28 days (in percentage points) in which the head, the spouse of the head and other working-age household members (aged 18-55) suffered from severe illness limiting daily activities. In addition, the regressions in columns 1, 2, 5 and 6 include county-by-year fixed effects and the set of covariates X discussed in Section 3; the regressions in columns 3, 4, 7 and 8 include county fixed effects and year fixed effects. Details of the variable construction and definitions are described in Appendix B.

Table A1: Auto-covariance Structure of the Health Shock

	Health shock
One-period lagged health shock	-0.663***
	(0.044)
Two-period lagged health shock	-0.313***
	(0.043)
R^2	0.293
N	3829

Note: This table reports the results from regression of the current health shock (defined for the head and the spouse of the head) on the first- and second-lagged health shock, and the county-by-year fixed effects. The county-by-year fixed effects are included such that the health shock is defined in a similar fashion as the main regression. The coefficients on the e lagged health shock terms are the estimated autocovariances of health shock at e lags. Standard errors are clustered at the county level.

Table A2: Impact of Health Shocks and Health Insurance: Health Shock Defined over the Past Three Months

Outcomes:	Regression Estimates α_0	α_1	$H_0 : \alpha_0 + \alpha_1 = 0$ (p-value)	Number of observations
A. Income				
Log (head and spouse income)	-0.006***	0.004	0.732	4122
	(0.001)	(0.006)		
Log (household income)	-0.002	-0.000	0.678	4590
	(0.001)	(0.006)		
B. Consumption				
Log (caloric intake)	-0.000	-0.001	0.214	4673
	(0.001)	(0.002)		
Log (protein intake)	0.001**	-0.001	0.193	4660
	(0.000)	(0.002)		
Log (fat intake)	-0.001	-0.002	0.168	4662
	(0.001)	(0.004)		
Log (carbohydrate intake)	-0.000	-0.002*	0.220	4658
	(0.000)	(0.001)		
C. Children's Outcomes				
% children in school	-0.007***	-0.000	0.040	2934
	(0.002)	(0.005)		
% children at work	0.007***	-0.005	0.654	3192
	(0.002)	(0.005)		
D. Agricultural Investment				
Log (investment in farming)	-0.002	-0.003	0.476	3491
	(0.002)	(0.007)		
Log (investement in livestock)	-0.008**	0.007	0.954	1762
	(0.004)	(0.012)		

*** p<0.01, ** p<0.05, * p<0.1. Each cell in regression columns 1 and 2 presents the coefficient estimates on health shock (α_0) and health shock interacted with the reform (α_1) in an OLS regression as specified in equation (3). Column 3 presents p-values of testing the null hypothesis that $\alpha_0 + \alpha_1 = 0$. Standard errors (in parentheses) are clustered at the county level. All dependent variables are first-differenced variables. Health shock is the change in the fraction of the past three months (in percentage points) in which the head and the spouse of the head suffered from severe illness limiting daily activities. Each regression includes the same set of covariates discussed in Section 3. Details of the variable construction and definitions are described in Appendix B.